**Computer Science Fundamentals**

1. How would you check if a linked list has cycles?

2. Given two elements in a binary search tree, find their lowest common ancestor.

3. Write a function to sort a given stack.

4. What is the time complexity of any comparison-based sorting algorithm? Can you prove it?

5. How will you find the shortest path from one node to another in a weighted graph? What if some weights are negative?

6. Find all palindromic substrings in a given string.

**Probability and Statistics**

1. The mean heights of men and women in a population were calculated to be μm and μw. What is the mean height of the total population?

2. A recent poll revealed that a third of the cars in Italy are Ferraris, and that half of those cars are red. If you spot a red car approaching from a distance, what is the likelihood that it is a Ferrari?

3. You’re trying to find the best place to put in an advertisement banner on your website. You can make the size (thickness) small, medium or large, and choose vertical position top, middle or bottom. At least how many total page visits (n) and ad clicks (m) do you need to say with 95% confidence that one of the designs performs better than all the other possibilities?

**Data Modeling and Evaluation**

1. A dairy farmer is trying to understand the factors that affect milk production of her cattle. She has been keeping logs of the daily temperature (usually 30-40°C), humidity (60-90%), feed consumption (2000-2500 kgs), and milk produced (500-1000 liters).

*a) How would you begin processing the data in order to model it, with the goal of predicting liters of milk produced in a day?*

*b) What kind of machine learning problem is this?*

2. Your company is building a facial expression coding system, which needs to take input images from a standard HD 1920×1080 pixel webcam, and continuously tell whether the user is in one of the following states: neutral, happy, sad, angry or afraid. When the user’s face is not visible in the camera frame, it should indicate a special state: none.

*a) What class of machine learning problems does this belong to?*

*b) If each pixel is made up of 3 values (for red, green, blue channels), what is the raw input data complexity (no. of dimensions) for processing each image? Is there a way to reduce the no. of dimensions?*

*c) How would you encode the output of the system? Explain why.*

3. Climate data collected over the past century reveals a cyclic pattern of rising and falling temperatures. How would you model this data (a sequence of average annual temperature values) to predict the average temperature over the next 5 years?

4. Your job at an online news service is to collect text reports from around the world, and present each story as a single article with content aggregated from different sources. How would you go about designing such a system? What ML techniques would you apply?

**Applying Machine Learning Algorithms and Libraries**

1. I’m trying to fit a single hidden layer neural network to a given dataset, and I find that the weights are oscillating a lot over training iterations (varying wildly, often swinging between positive and negative values). What parameter do I need to tune to address this issue?

2. When training a support vector machine, what value are you optimizing for?

3. Lasso regression uses the L1-norm of coefficients as a penalty term, while ridge regression uses the L2-norm. Which of these regularization methods is more likely to result in *sparse* solutions, where one or more coefficients are exactly zero?

4. When training a 10-layer neural net using backpropagation, I find that the weights for the top 3 layers are not changing at all! The next few layers (4-6) are changing, but very slowly. What’s going on and how do I fix this?

5. I’ve found some data about wheat-growing regions in Europe that includes annual rainfall (R, in inches), mean altitude (A, in meters) and wheat output (O, in kgs/km2). A rough analysis and some plots make me believe that output is related to the square of rainfall, and log of altitude: O = *β*0 + *β*1 × R2 + *β*2 × loge(A)  
Can I fit the coefficients (*β*) in my model to the data using linear regression?

**Software Engineering and System Design**

1. You run an ecommerce website. When a user clicks on an item to open its details page, you would like to suggest 5 more items that the user may be interested in, based on item features as well as the user’s purchase history, and display them at the bottom of the page. What services and database tables would you need to support this behavior? Assuming they’re available, write a query or procedure to fetch the 5 items to suggest.

2. What data would you like to collect from an online video player (like YouTube) to measure user engagement and video popularity?

3. A very simple spam detection system works as follows: It processes one email at a time and counts the number of occurrences of each unique word in it (term frequency), and then it compares those counts with those of previously seen emails which have been marked as spam or not. In order to scale up this system to handle a large volume of email traffic, can you design a map-reduce scheme that can run on a cluster of computers?

4. You want to generate a live visualization of what portion of a webpage users are currently viewing and clicking, sort of like a heat map. What components/services/APIs do you need in place, on the client and server end, to enable this?

**Computer Vision Focus**

1. How could you derive feature from image data?

2. Is drawing a bounding box over a region of interest a classification or regression problem (or both)? Explain

3. You are given a set of images and their extracted features (say histogram of oriented gradients). Describe how you would make an object detection system for 10 different objects. Assume the entire frame of the image is taken up by the object. i.e. no bounding boxes need to be drawn.

4. You are now given a new set of images that belong to an 11th class. How can you adjust your model to account for this new class?

5. Why does deep learning work so well in computer vision?

6. Describe a way to organize set of unlabeled images?

**NLP Focus**

1. How can you design a parts-of-speech (POS) tagger?

2. How would you deal with unseen words at test time?

3. Describe a way to extract keywords from a set of documents.

4. How can you build a system to classify documents as spam or not spam?

5. What kind of techniques can you use to compare the similarity of two documents?

6. What is named entity recognition? How can you train your own model?

7. What is entropy in the field of information theory?